





603	612	621	630	639	648
ATG TCT TCC CCA GCG CTG ATG GCA GGT GCT TCA GGA AAG GTG ACT GAC TTC AAC					
M S S P A L M A G A S G K V T D F N					
657	666	675	684	693	702
AAC GGC ACC TAC CTG GTC AGC TTC ACT CTG TTC TGG GAG GGC GTG TCT CTG					
N G T Y L V S F T L F W E G Q V S L					
711	720	729	738	747	756
TCT CTG CTG CTC ATC CAC CAC CCC AGT GAA GGG GTG TCA GCT CTC TGG AGT GCA AGG					
S L L L I H P S E G V S A L W S A R					
765	774	783	792	801	810
AAC CAA GGC TAT GAC AGG GTG ATC TTC ACT GGC CAG TTT GTG AAT GGC ACT TCC					
N Q G Y D R V I F T G Q F V N G T S					
819	828	837	846	855	864
CAA GTC CAC TCT GAA TGT GGC CTG ATC CTA AAC ACA AAT GCT GAA TTG TGC CAG					
Q V H S E C G L I L N T N A E L C Q					

**FIGURE 1C**

873	882	891	900	909	918
TAC CTG GAC AAC AGA GAC CAA GAA GGC TTC TAC TGT GTG AGG CCT CAA CAC ATG					
---	---	---	---	---	---
Y L D N R D Q E G F Y C V R P Q H M					
927	936	945	954	963	972
CCC TGT GCT GCA CTC ACT CAC ATG TAT TCT AAG AAC AAG AAA GTT TCT TAT CTT					
---	---	---	---	---	---
P C A A L T H M Y S K N K K V S Y L					
981	990	999	1008	1017	1026
AGC AAA CAA GAA AAG AGC CTC TTT GAA AGG TCA AAT GTG GGT GTA GAG ATT ATG					
---	---	---	---	---	---
S K Q E K S L F E R S N V G V E I M					
1035	1044	1053	1062	1071	1080
GAA AAA TTC AAT ACA ATT AGT GTC TCC AAA TGC AAC ACA CTG AAG TCA GTG GAT					
---	---	---	---	---	---
E K F N T I S V S K C N T L K S V D					
1089	1098	1107	1116	1125	1134
CTG CAT GAA TCT GGA AAA TTG CAA CAC CAG CTT GCT GTG GAT TTG GAT AGG AAC					
---	---	---	---	---	---
L H E S G K L Q Q H Q Q L A V D L D R N					

**FIGURE 1D**

1143	1152	1161	1170	1179	1188
ATC AAC ATC CAG TGG CAA AAA TAT TGT TAT CCC TTG ATA GGA TCA ATG ACC TAT					
---	---	---	---	---	---
I N I Q W Q K Y C Y P L I G S M T Y					
1197	1206	1215	1224	1233	1242
TCA GTC AAA GAG ATG GAG TAC CTC ACC CGG GCC ATT GAC AGA ACT GGA GGA GAA					
---	---	---	---	---	---
S V K E M E Y L T R A I D R T G G E					
1251	1260	1269	1278	1287	1296
AAA AAT ACT GTC ATT GTT ATT TCC CTG GGC CAG CAT TTC AGA CCC TTT CCC ATT					
---	---	---	---	---	---
K N T V I V I S L G Q H F R P F I					
1305	1314	1323	1332	1341	1350
GAT GTT TTT ATC CGA AGG GCC CTC AAT GTC CAC AAA GCC ATT CAG CAT CTT CTT					
---	---	---	---	---	---
D V F I R R A L N V H K A I Q H L L					
1359	1368	1377	1386	1395	1404
CTG AGA AGC CCA GAC ACT ATG GTT ATC ATC AAA ACA GAA AAC ATC AGG GAG ATG					
---	---	---	---	---	---
L R S P D T M V I I K T E N I R E M					

**FIGURE 1E**

1413	1422	1431	1440	1449	1458
TAC AAT GAT GCA GAA AGA TTT AGT GAC TTT CAT GGT TAC ATT CAA TAT CTC ATC					
Y N D A E R F S D F H G Y I Q Y L I					
1467	1476	1485	1494	1503	1512
ATA AAG GAC ATT TTC CAG GAT CTC AGT GTG AGT ATC ATT GAT GCC TGG GAT ATA					
I K D I F Q D L S V S I I D A W D I					
1521	1530	1539	1548	1557	1566
ACA ATT GCA TAT GGC ACA AAT AAT GTA CAC CCA CCT CAA CAT GTA GTC GGA AAT					
T I A Y G T N N V H P P Q H V V G N					
1575	1584	1593	1602	1611	
CAG ATT AAT ATA TTA AAC TAT ATT TGT TAA ATA ACA AAA AAA AAA A 3'					
Q I N I L L N Y I C					

**FIGURE 1F**

5'	9	18	27	36	45	54
GCC	AGT TAA GAA CTC ACA GGT ACT TTT CCA GCC ACC CAG ATA GGA GAG ATC ATT					
63	72	81	90	99	108	
AAA	ACA GTG CAT TCT GTG CTA CCT GAC ACC TAT TGG GGT CCT GGA AGG AAG					
117	126	135	144	153	162	
CAA	CAA TCC TGA GTG AAA CCT CGA CAA GAA GTA TCC AAT AGG ACA TTC GTC ATG					
					---	
					M	
171	180	189	198	207	216	
TCC	TCA AAT ACA ATG CTT CAA AAA ACG CTG CTG ATC TTG ATC TTT TCA GTA					
					---	
S	S N T M L L Q K T L L I L I S F S V					
225	234	243	252	261	270	
GTA	ACC TGG ATG ATT TTT ATA ATT TCT CAG AAC TTC ACA AAG CTT TGG TCT GCT					
					---	
V	T W M I F I I S Q N F T K L W S A					
279	288	297	306	315	324	
CTA	AAC TTA TCC ATC TCT GTC CAT TAC TGG AAC AAC TCC GCA AAG TCC TTA TTC					
					---	
L	N L S I S V H Y W N N S A K S L F					

**FIGURE 2A**

333	342	351	360	369	378
CCT AAA ACA TCA CTG ATA CCA TTA AAG CCA CTA ACA GAG ACT GAA CTC AGA ATA					
---	---	---	---	---	---
P K T S L I P L K P L T E T E L R I					
387	396	405	414	423	432
AAG GAA ATC ATA GAG AAA CTA GAT CAG CAG ATC CCA CCC AGA CCT TTC ACC CAT					
---	---	---	---	---	---
K E I I E K L D Q Q Q I P P P R P F T H					
441	450	459	468	477	486
GTG AAC ACC ACC AGT GCC ACA CAC CAC AGC ACC ACC ATC CTC AAC CCT CGA					
---	---	---	---	---	---
V N T T T S A T H S T A T I L N P R					
495	504	513	522	531	540
GAT ACA TAC TGC AGG GGA GAC CAG CTG GAC ATC CTA CTG GAG GTG AGG GAC CAC					
---	---	---	---	---	---
D T Y C R G D Q L D I L L E V R D H					
549	558	567	576	585	594
TTG GGA CAG AGG AAG CAA TAT GGT GGG GAT TTC CTG AGG GCC AGG ATG TCC TCC					
---	---	---	---	---	---
L G Q R K Q Y G G D D F L R A R M S S					

**FIGURE 2B**



603	612	621	630	639	648
CCA GCA CTG	ACG GCA GGT	GCT TCA GGA AAG	GTG ATG GAC	TTC AAC AAT	GGC ACC
---	---	---	---	---	---
P A L T A G A S G K V M D F N N G T					
657	666	675	684	693	702
TAC CTG GTC	AGC TTC ACT	CTG TTC TGG GAG	GGC CAG GTC	TCC CTG TCT	CTG CTG
---	---	---	---	---	---
Y L V S F T L F W E G Q V S L S L L					
711	720	729	738	747	756
CTC ATC CAC	CCC AGT GAA	GGG GCG TCG	GCT CTC TGG AGG	GCA AGG AAC	CAA GGC
---	---	---	---	---	---
L I H P S E G A S A L W R A R N Q G					
765	774	783	792	801	810
TAT GAT AAA	ATT ATT TTC	AAA GGC AAA	TTT GTT AAT	GGC ACC	TCT CAT
---	---	---	---	---	---
Y D K I I F K G K F V N G T S H V F					
819	828	837	846	855	864
ACT GAA TGT	GGC CTG ACC	CTA AAC TCA	AAT GCT GAA	CTC TGT	GAA TAT
---	---	---	---	---	---
T E C G L T L N S N A E L C E Y L D					

FIGURE 2C

873	882	891	900	909	918
GAC AGA GAC CAA GAA GCC TTC TAT TGT ATG AAG CCT CAA CAC ATG CCC TGT GAG					
D R D Q E A F Y C M K P Q H M P C E					
927	936	945	954	963	972
GCT CTG ACC TAC ATG ACC ACC CGG AAT AGA GAG GTA TCT TAT CTT ACA GAC AAG					
A L T Y M T T R N R E V S Y L T D K					
981	990	999	1008	1017	1026
GAA AAC AGC CTT TTC CAC AGG TCC AAA GTG GGA GTT GAA ATG ATG AAG GAT CGT					
E N S L F H R S K K V G V E M M K D R					
1035	1044	1053	1062	1071	1080
AAA CAC ATT GAT GTC ACT AAT TGT AAC AAG AGA GAA AAA ATA GAA GAG ACA TGC					
K H I D V T N C N K R E K I E E T C					
1089	1098	1107	1116	1125	1134
CAA GTT GGA ATG AAG CCT CCT GTC CCT GGT GGT TAT ACT TTA CAA GGA AAA TGG					
Q V G M K P P V P P G G Y T L Q G K W					

FIGURE 2D

1143	1152	1161	1170	1179	1188
ATA ACA ACA TTT TGC AAC CAG GTT CAG TTA GAC ACA ATT AAG ATA AAT GGC TGT					
---	---	---	---	---	---
I T T F C N Q V Q L D T I K I N G C					
1197	1206	1215	1224	1233	1242
TTG AAA GGC AAA CTC ATT TAC CTC CTG GGA GAC TCT ACA CTA CGT CAG TGG ATC					
---	---	---	---	---	---
L K G K L I Y L L G D S T L R Q W I					
1251	1260	1269	1278	1287	1296
TAC TAC TTC CCC AAA GTT GTA AAA ACA CTG AAG TTT TTT GAT CTT CAT GAA ACT					
---	---	---	---	---	---
Y Y F P K V V K T L L K F F D L H E T					
1305	1314	1323	1332	1341	1350
GGA ATC TTT AAG AAA CAT TTG CTT CTG GAT GCA GAA AGA CAC ACT CAG ATT CAA					
---	---	---	---	---	---
G I F K K H L L L L D A E R H T Q I Q					
1359	1368	1377	1386	1395	1404
TGG AAA AAA CAT AGC TAT CCC TTC GTC ACT TTC CAG CTC TAC TCT CTG ATA GAT					
---	---	---	---	---	---
W K K H S Y P F V T T F Q L Y S L I D					

**FIGURE 2E**

1413	1422	1431	1440	1449	1458
CAT GAT TAT ATC CCT CGG GAA ATT GAC CGG CTA TCA GGT GAC AAA AAC ACA GCC					
---	---	---	---	---	---
H D Y I P R E I D R L S G D K N T A					
1467	1476	1485	1494	1503	1512
ATC GTC ATC ACC TTT GGC CAG CAC TTT AGA CCA TTT CCC ATT GAC ATT TTT ATT					
---	---	---	---	---	---
I V I T F G Q H F R P F P I D I F I					
1521	1530	1539	1548	1557	1566
CGC AGG GCC ATC GGT GTT CAA AAG GCT ATT GAA AGA CTG TTC CTA AGA AGC CCA					
---	---	---	---	---	---
R R A I G V Q K A I E R L F L R S P					
1575	1584	1593	1602	1611	1620
GCC ACT AAA GTG ATT ATT AAG ACA GAA AAC ATC AGG GAG ATG CAC ATA GAG ACA					
---	---	---	---	---	---
A T K V I I K T E N I R E M H I E T					
1629	1638	1647	1656	1665	1674
GAG AGG TTT GGA GAC TTC CAT GGT TAT ATT CAC TAT CTT ATC ATG AAG GAT ATT					
---	---	---	---	---	---
E R F G D F H G Y I H Y L I M K D I					

FIGURE 2F

1683	1692	1701	1710	1719	1728
TTC AAA GAC CTC AAC GTG GGC ATC ATT GAT GCC TGG GAC ATG ACC ATT GCA TAT					
---	---	---	---	---	---
F K D L N V G I I D A W D M T I A Y					
1737	1746	1755	1764	1773	1782
GGC ACT GAC ACT ATC CAC CCA CCT GAT CAT GTG ATT GGA AAT CAG ATT AAC ATG					
---	---	---	---	---	---
G T D T I H P P D H V I G N Q I N M					
1791	1800	1809	1818	1827	1836
TTC TTA AAC TAC ATT TGC TAA GGG ATA AAT ACT ATA CAA AAT CAC TAG GAA CCA					
---	---	---	---	---	---
F L N Y I C					
1845	1854	1863	1872	1881	1890
ATC TCT GCA CAT AAT CCC ACA TGT ATT GTA AAG TAA GTT TTA CTC ATT TTA GGA					
1899	1908	1917	1926	1935	1944
ACT AAG GAA AAT AAA TTT AAA AGA ATC TGT TTG GGG AGG AAG GCT ATG TAA GGA					
1953	1962	1971	1980	1989	1998
CAA TGA CAA CTG ATA AGG GAT GCA AAA CCA AGA GAA TCA TTC ATG AAG AAT GAC					
2007	2016	2025	2034	2043	2052
TAT ACC ATG CCT GGT TCT GAT GCT CGT TTA AAA TAT TAA AAA AGT TTT TTA AAA					

**FIGURE 2G**

2061	2070	2079	2088	2097	2106
GCC ATG TTA TTA AGC TGA TTT GAA AAT ATC TGT ACA AAT TCA TGA TGC TTT CTA					
2115	2124	2133	2142	2151	2160
TTT CCA ATA TAG ATA TTT CCT AGC TCT GTC TAT TGA AAA GGC CTA GGA GCA ATG					
2169	2178	2187	2196	2205	2214
ATA ACC CAT TAG CAA TAA TCA CTC CGA GCA CCC TAA CTG TGA TGT CTA AGA ACC					
2223	2232	2241			
CTT CCT CAA TAA AAG AAA AGA GGC ATC CTT GAA G 3'					

**FIGURE 2H**

1	M	K	I	S	M	I	N	Y	K	S	L	L	A	L	L	-	F	I	L	A	S	W	I	I	F	T	V	F	Q	N	3229449
1	M	S	S	N	T	M	L	Q	K	T	L	L	I	L	I	S	F	S	V	T	W	M	I	I	F	I	I	S	Q	N	7484349
1	M	-	-	-	-	L	H	K	-	Y	L	K	L	I	C	L	L	A	A	I	C	V	L	C	I	I	S	Q	N	g1762	
30	S	T	K	V	W	S	A	L	N	L	S	I	S	L	H	Y	W	N	N	S	T	K	S	L	F	P	K	T	P	L	3229449
31	F	T	K	L	W	S	A	L	N	L	S	I	S	V	H	Y	W	N	N	S	A	K	S	L	F	P	K	T	S	L	7484349
25	S	T	K	I	W	G	A	L	K	L	P	N	S	H	Y	Y	S	N	T	S	M	I	S	I	P	K	M	S	V	g1762	
60	I	S	L	K	P	L	T	E	T	E	L	R	I	K	E	I	I	E	K	L	D	Q	Q	I	P	P	R	P	F	T	3229449
61	I	P	L	K	P	L	T	E	T	E	L	R	I	K	E	I	I	E	K	L	D	Q	Q	I	P	P	R	P	F	T	7484349
55	S	P	V	K	S	L	T	E	T	E	L	R	V	K	E	I	L	E	K	L	D	R	L	I	P	P	R	P	F	T	g1762
90	H	V	N	T	T	S	A	T	H	S	T	A	T	I	L	N	P	R	D	T	Y	C	R	G	D	Q	L	H	I	3229449	
91	H	V	N	T	T	S	A	T	H	S	T	A	T	I	L	N	P	R	D	T	Y	C	R	G	D	Q	L	D	I	7484349	
85	H	V	N	T	T	S	A	T	H	S	T	A	T	I	L	N	P	Q	D	K	Y	C	V	G	D	Q	L	D	I	g1762	
120	L	L	E	V	R	D	H	L	G	R	R	K	Q	Y	G	G	D	F	L	R	A	R	M	S	S	P	A	L	M	A	3229449
121	L	L	E	V	R	D	H	L	G	Q	R	K	Q	Y	G	G	D	F	L	R	A	R	M	S	S	P	A	L	T	A	7484349
115	L	L	E	V	R	D	Y	L	G	H	Q	K	E	Y	G	G	D	F	L	R	A	R	M	F	S	P	A	L	K	A	g1762
150	G	A	S	G	K	V	T	D	F	N	N	G	T	Y	L	V	S	F	T	L	F	W	E	G	Q	V	S	L	S	L	3229449
151	G	A	S	G	K	V	M	D	F	N	N	G	T	Y	L	V	S	F	T	L	F	W	E	G	Q	V	S	L	S	L	7484349
145	G	A	S	G	K	V	T	D	F	N	N	G	T	Y	L	V	S	F	T	L	F	W	E	G	Q	V	S	L	S	V	g1762

FIGURE 3A

180	L	L	I	H	P	S	E	G	V	S	A	L	W	S	A	R	N	Q	G	Y	D	R	V	I	F	T	G	Q	F	V	3229449
181	L	L	I	H	P	S	E	G	A	S	A	L	W	R	A	R	N	Q	G	Y	D	K	I	I	F	K	G	K	F	V	7484349
175	L	L	I	H	P	S	E	G	A	S	A	L	W	R	A	R	N	Q	G	Y	D	R	I	I	F	K	G	Q	F	V	g1762
210	N	G	T	S	Q	V	H	S	E	C	G	L	I	L	N	T	N	A	E	L	C	Q	Y	L	D	N	R	D	Q	E	3229449
211	N	G	T	S	H	V	F	T	E	C	G	L	T	L	N	S	N	A	E	L	C	E	Y	L	D	D	R	D	Q	E	7484349
205	N	G	T	S	H	V	F	T	E	C	S	L	T	L	N	S	N	T	E	E	C	K	Y	L	N	G	R	D	Q	D	g1762
240	G	F	Y	C	V	R	P	Q	H	M	P	C	A	A	L	T	H	M	Y	S	K	N	K	K	V	S	Y	L	S	K	3229449
241	A	F	Y	C	M	K	P	Q	H	M	P	C	E	A	L	T	Y	M	T	T	R	N	R	E	V	S	Y	L	T	D	7484349
235	V	F	Y	C	M	K	P	Q	H	M	P	C	E	A	L	T	H	V	T	S	R	N	R	D	I	S	Y	L	T	S	g1762
270	Q	E	K	S	L	F	E	R	S	N	V	G	V	E	I	M	E	K	F	N	T	I	S	V	S	K	C	N	-	-	3229449
271	K	E	N	S	L	F	H	R	S	K	V	G	V	E	M	M	K	D	R	K	H	I	D	V	T	N	C	N	K	R	7484349
265	K	E	K	N	L	F	H	R	S	K	V	G	V	E	I	M	K	N	Q	-	H	I	D	V	S	Q	C	N	K	S	g1762
298	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3229449
301	E	K	I	E	T	C	Q	V	G	M	K	P	P	V	P	G	G	Y	T	L	Q	G	K	W	I	T	T	F	C	7484349	
294	K	E	V	K	E	C	Q	I	G	M	K	I	P	V	P	G	G	Y	T	L	Q	G	R	W	L	T	T	F	C	g1762	
298	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3229449
331	N	Q	V	Q	L	D	T	I	K	I	N	G	C	L	K	G	K	L	I	Y	L	L	G	D	S	T	L	R	Q	W	7484349
324	N	Q	I	Q	L	D	T	A	K	I	S	G	C	L	K	G	K	L	I	Y	L	M	G	D	S	T	L	R	Q	W	g1762

FIGURE 3B



298	- - - - -	T L K S V D L H E S G K L Q H Q L A V D L	3229449
361	I Y Y F P K V V K T L K F F D L H E T G I F K K H L L L D A	7484349	
354	I Y Y L P K V M K T L K F F D L H E T G N F K K H L L L D A	g1762	
319	D R N I N I Q W Q K Y C Y P L I G S M T Y S V K E M E Y L T	3229449	
391	E R H T Q I Q W K K H S Y P F V T F Q L Y S L I D H D Y I P	7484349	
384	E K H T Q I Q W K K H S H P F V T Y Q L F S V I D H G Y I P	g1762	
349	R A I D R T G G E K N T V I V I S L G Q H F R P F I D V F	3229449	
421	R E I D R L S G D K N T A I V I T F G Q H F R P F I D I F	7484349	
414	Q E I D R L I G D K D T V I V I T F G Q H F R P F I D I F	g1762	
379	I R R A L N V H K A I Q H L L R S P D T M V I I K T E N I	3229449	
451	I R R A I G V Q K A I E R L F L R S P A T K V I I K T E N I	7484349	
444	I R R A I S V R Q A I E R L F L R S P A T K V I V K T E N I	g1762	
409	R E M Y N D A E R F S D F H G Y I Q Y L I I K D I F Q D L S	3229449	
481	R E M H I E T E R F G D F H G Y I H Y L I M K D I F K D L N	7484349	
474	R E M H I E A E R F G D F H G Y I Q Y L T L K D I F K D L N	g1762	
439	V S I I D A W D I T I A Y G T N N V H P P Q H V V G N Q I N	3229449	
511	V G I I D A W D M T I A Y G T D T I H P P D H V I G N Q I N	7484349	
504	V G V V D A W D M T I A Y G T N N V H P P D Q V I G N Q I N	g1762	

FIGURE 3C

bioRxiv preprint doi: <https://doi.org/10.1101/2017.04.11.128444>; this version posted April 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3229449  
7484349  
g1762

469	I	L	L	N	Y	I	C
541	M	F	L	N	Y	I	C
534	M	F	L	N	Y	I	C

FIGURE 3D